ABSTRACT

Temporal evolution of data assimilated solutions are physically inconsistent when corresponding model errors are not determined. The smoother can be identified as estimating model errors by inverting the data increments. A novel RTS smoother is introduced that approximates the model error evaluation by partitioning it into smaller independent elements. The approximation renders high-resolution global eddy-resolving data assimilation computationally feasible. The importance of identifying effective representations of model error sources is emphasized.

"Physical Consistency of Data Assimilated Estimates and the ECCO Hierarchical Assimilation Method "

I.Fukumori, T.Lee, D.Menemenlis, L.-L. Fu, and the ECCO Group

Jet Propulsion Laboratory, California Institute of Technology

Consortium for Estimating the Circulation and Climate of the Ocean (ECCO/JPL-MIT-SIO)

Physical Consistency

Temporal evolution of most data assimilated solutions, including results of the Kalman filter, are not physically consistent; e.g., changes in energy and mass cannot fully be accounted for. Such inconsistencies are often ignored or forgotten, but the deficiencies are important especially when data assimilated solutions are used to understand mechanisms of ocean circulation, such as heat and tracer balances [3, 4]. The problem is illustrated by examples in Figures 1, 2, 3, and 4. The mathematical problem and its solution are described in the next section.

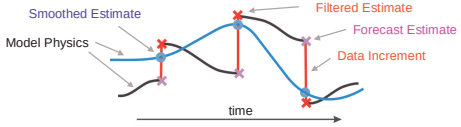

Fig 1: Schematic of Model State Evolution (e.g., Temperature). *Most sequential assimilation methods integrate the model in time (**black**) and correct the state according to observations when available (**red**). The **black** model evolution is physically accounted for, such as by advection, mixing, and external forcing. However, the **red** data corrections are not identified by any particular process, but are supposed to reflect implicit combinations of the effects of various model errors (errors in advection, mixing, and forcing). The **blue** smoothed estimate inverts the **red** data corrections and explicitly corrects the state and model errors, that result in a state evolution that can be dynamically and physically accounted for (**blue**).*

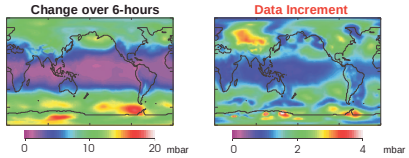

Fig 2: Atmospheric Mass Budget (NCEP re-analysis). *Standard deviation of surface pressure for **6-hour forecast (left)** and **6-hourly data increments (right).** On average 24% of mass change every 6-hours is not physically accounted for in the NCEP re-analyses.*

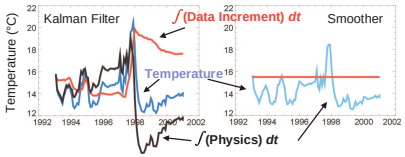

Fig 3: Physical Consistency of Model Temperature Evolution. *Temperature budget in the tropical Pacific Ocean (100-150m, 5°S~5°N, 150°W~90°W) based on a data assimilated estimate (ECCO-2, [2]). Time-integration (sum) of Kalman filtering **data increments** is as large as the effects of **model physics** (left). In comparison, temperature of the **smoothed estimate** (right) is consistent with **model physics** (black and blue curves are coincident).*

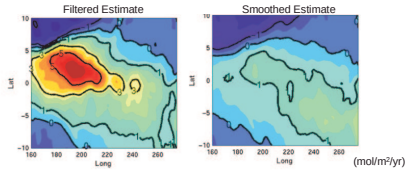

Fig 4: CO2 Sea-Air Flux During the '97-'98 El Nino Derived from ECCO-2 [4]. *The **filtered solution's** (left) unrealistically large CO2 flux during ENSO (due to inconsistencies in circulation) is corrected in the **smoothed estimate (right).***

The Mathematical Problem and Its Solution

The data assimilation problem can be identified as an inverse problem of solving a set of simultaneous equations for the model state (vector $\mathbf{x}$) and model errors (control vector $\mathbf{u}$), as a function of time, given observations (vector $\mathbf{y}$) and the theoretical relationships among the elements ($\mathbf{H}$, $\mathbf{A}$, $\mathbf{G}$) (Eq 1).

Observations → $\mathbf{H}\mathbf{x}_t$

Model Eqs → $\mathbf{x}_{t+1} - \mathbf{A}\mathbf{x}_t - \mathbf{G}\mathbf{u}_t$

$$= \begin{pmatrix} \vdots \\ \mathbf{y}_t \\ \vdots \\ \mathbf{0} \\ \vdots \end{pmatrix} \qquad (1)$$

$\mathbf{x}$: model state
$\mathbf{y}$: observations
$\mathbf{H}$: observation operator
$\mathbf{A}$, $\mathbf{G}$: model dynamics
$\mathbf{u}$: control (e.g., forcing, and model errors)

The least-squares solution to a generic (linear) inverse problem $\mathbf{E}\mathbf{a} = \mathbf{b}$ is;

$$\hat{\mathbf{a}} = \mathbf{a}_0 + \mathbf{R}_{aa}\mathbf{E}^T\left(\mathbf{E}\mathbf{R}_{aa}\mathbf{E}^T + \mathbf{R}_{bb}\right)^{-1}\left(\mathbf{b} - \mathbf{E}\mathbf{a}_0\right) \qquad (2)$$

where, $\mathbf{a}_0$ is first guess and $\mathbf{R}_{aa}$, $\mathbf{R}_{bb}$ are error covariance matrices of $\mathbf{a}_0$ and $\mathbf{b}$

The **Kalman Filter** can be recognized as inverting the observations part of the simultaneous set of equations (1). The first guess given by the model,

$$\tilde{\mathbf{x}}_t = \mathbf{A}\hat{\mathbf{x}}_{t-1} + \mathbf{G}\mathbf{u}_{t-1} \qquad (3)$$

is corrected by the observations by

$$\hat{\mathbf{x}}_t = \tilde{\mathbf{x}}_t + \tilde{\mathbf{P}}_t\mathbf{H}^T\left(\mathbf{H}\tilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R}_t\right)^{-1}\left(\mathbf{y}_t - \mathbf{H}\tilde{\mathbf{x}}_t\right) \qquad (4)$$

where, $\tilde{\mathbf{P}}_t$, $\mathbf{R}_t$ are error covariance matrices of $\tilde{\mathbf{x}}_t$ and $\mathbf{y}_t$. Note the correspondence between Eqs (2) and (4). Combining (3) and (4) describes the evolution of the filtered solution;

$$\hat{\mathbf{x}}_t = \mathbf{A}\hat{\mathbf{x}}_{t-1} + \mathbf{G}\mathbf{u}_{t-1} + \tilde{\mathbf{P}}_t\mathbf{H}^T\left(\mathbf{H}\tilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R}_t\right)^{-1}\left(\mathbf{y}_t - \mathbf{H}\tilde{\mathbf{x}}_t\right) \qquad (5)$$

The first two terms represent the model physics (black curve in Fig 1). The last term is the data increment (red line in Fig 1) that represents errors in the first two terms; however, which of the two or what component of the two the last term corresponds to is not determined.

The **Rauch-Tung-Striebel Fixed-Interval Smoother** can be identified as projecting the last term in (5) to the elements of the first two by inverting the model equations using the filtered solution as a first guess. After filtering is performed to the end, $\bar{\mathbf{x}}_T = \hat{\mathbf{x}}_T$ we look for a new solution $(\mathbf{x}_t, \mathbf{u}_t)$ backwards in time that satisfies the model equations;

$$\bar{\mathbf{x}}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{G}\mathbf{u}_{t-1} = \begin{pmatrix} \mathbf{A} & \mathbf{G} \end{pmatrix}\begin{pmatrix} \mathbf{x}_{t-1} \\ \mathbf{u}_{t-1} \end{pmatrix} \qquad (6)$$

Eq (6) is merely another inverse problem, and can be solved using the filtered solution as the first guess as in Eq (2). The correspondence with Eq (2) is

$$\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{G} \end{pmatrix}, \quad \mathbf{b} = \bar{\mathbf{x}}_t, \quad \mathbf{a}_0 = \begin{pmatrix} \hat{\mathbf{x}}_{t-1} \\ \mathbf{u}_{t-1} \end{pmatrix}$$

$$\mathbf{R}_{nn} = 0, \quad \mathbf{R}_{aa} = \begin{pmatrix} \mathbf{P}_{t-1} & 0 \\ 0 & \mathbf{Q}_{t-1} \end{pmatrix} \qquad (7)$$

where $\mathbf{Q}$ is the error covariance matrix of model error source (process noise) $\mathbf{u}$. By substituting (7) to (2), we have the familiar RTS smoother;

$$\begin{pmatrix} \bar{\mathbf{x}}_{t-1} \\ \bar{\mathbf{u}}_{t-1} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{x}}_{t-1} \\ \mathbf{u}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{P}_{t-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T + \mathbf{G}\mathbf{Q}_{t-1}\mathbf{G}^T\right)^{-1} \\ \mathbf{Q}_{t-1}\mathbf{G}^T\left(\mathbf{G}\mathbf{Q}_{t-1}\mathbf{G}^T + \mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T\right)^{-1} \end{pmatrix}\left(\bar{\mathbf{x}}_t - \mathbf{A}\hat{\mathbf{x}}_{t-1} - \mathbf{G}\mathbf{u}_{t-1}\right) \qquad (8)$$

Note, again, the correspondence between Eqs (8) and (2). From (8) and its derivation, the smoothed solution $(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ is shown to satisfy the model physics (model constraints; e.g., continuity). ***This illustrates the importance of explicitly modeling (identifying) model error sources by physically sensible processes ($\mathbf{G}$ and $\mathbf{Q}$).*** The last expression in parenthesis can be rewritten as,

$$\bar{\mathbf{x}}_t - \mathbf{A}\hat{\mathbf{x}}_{t-1} - \mathbf{G}\mathbf{u}_{t-1} = \bar{\mathbf{x}}_t - \tilde{\mathbf{x}}_t = \left(\bar{\mathbf{x}}_t - \hat{\mathbf{x}}_t\right) + \left(\hat{\mathbf{x}}_t - \tilde{\mathbf{x}}_t\right) \qquad (9)$$

Namely, the smoother (Eq 8) projects (inverts) the sum of the smoother correction (increment by formally future data) and filter correction (data increment) to the model state and to the model errors (control), backwards in time. Typical formulation of adjoint methods [3, 6] and Green's function methods [5] provide alternate direct approaches to solving for the controls.

The Partitioned Kalman Filter (PKF) and Smoother (PS) [1]

The central difficulty of Kalman filtering is the computational evaluation of the time-evolving state error covariance matrix $\mathbf{P}_t$ that requires a number of operations proportional to the cube of the matrix's dimension. Then, a significant computational savings can be achieved by approximating $\mathbf{P}_t$ by a sum of smaller independent elements that are evaluated separately from each other and then combined together. For instance, small-scale errors in the North Atlantic may be evaluated separately from those of the North Pacific, etc. If the state (dimension N) is partitioned into L equal elements, each partition will require operations proportional to $(N/L)^3$ to evaluate the error covariance. Combined together with other partitions, the overall operation count will be proportional to N^3/L^2 *instead of* N^3, which is a savings by a factor L^2.

Mathematically, the model state error $\delta\mathbf{x}$ is approximated by a sum of L statistically independent elements $\delta\mathbf{x}'_i$, each with smaller dimension;

$$\delta\mathbf{x} \approx \mathbf{B}_1\delta\mathbf{x}'_1 + \cdots + \mathbf{B}_L\delta\mathbf{x}'_L \approx \sum_i^L \mathbf{B}_i\delta\mathbf{x}'_i \qquad (10)$$

where $\dim(\delta\mathbf{x}'_i) << \dim(\delta\mathbf{x})$ and $\mathbf{B}_i$ defines a transformation operator (e.g.,Figs 5 and 6). Given their independence, the overall state error covariance matrix can be approximated by those of the individual partitions as

$$\mathbf{P} \approx \sum_i^L \mathbf{B}_i\mathbf{P}'_i\mathbf{B}_i^T \qquad (11)$$

where $\mathbf{P}'_i \equiv \left\langle \delta\mathbf{x}'_i\delta\mathbf{x}'^T_i \right\rangle$. The smaller dimension of $\mathbf{P}'_i$ relative to that of $\mathbf{P}_i$ allows for the significant computational savings in approximating $\mathbf{P}_i$ by (11). By substituting Eq (11) into definitions of the Kalman filter and smoother, the overall filter and smoother gains can be approximated by the sum of equivalent operators of the individual partitions;

$$\mathbf{K} \approx \sum_i^L \mathbf{B}_i\mathbf{K}'_i, \quad \mathbf{K}'_i \equiv \hat{\mathbf{P}}'_i\mathbf{H}_i'^T\mathbf{R}^{-1}$$

$$\mathbf{S} \approx \sum_i^L \mathbf{B}_i\mathbf{S}'_i\mathbf{B}_i^*, \quad \mathbf{S}'_i \equiv \hat{\mathbf{P}}'_i\mathbf{A}_i'^T\tilde{\mathbf{P}}_i'^{-1} \qquad (12)$$

Likewise, the filter and smoother increments can be approximated by sums of equivalent increments of separate partitions.

$$\delta_K \equiv \hat{\mathbf{x}} - \tilde{\mathbf{x}} \equiv \mathbf{K}\left(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}\right) \approx \sum_i \mathbf{B}_i\delta'_K, \quad \delta'_K \equiv \hat{\mathbf{x}}' - \tilde{\mathbf{x}}' = \mathbf{K}'_i\left(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}\right)$$

$$\delta_S \equiv \bar{\mathbf{x}} - \hat{\mathbf{x}} \approx \sum_i \mathbf{B}_i\delta'_S, \quad \delta'_S(t) \equiv \bar{\mathbf{x}}' - \hat{\mathbf{x}}' = \mathbf{S}'_i\left(\delta'_S(t+1) + \delta'_K(t+1)\right) \qquad (13)$$

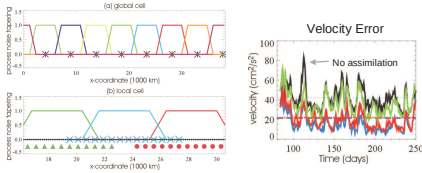

Fig 5: Example of Partitioning (left) and Skill of Partitioned Smoother (right) . *A PKF/PS is constructed for a 1-dimensional non-rotating shallow water model in [1]. The partitioning consists of a global coarse cell (a) and overlapping local fine cells (b) that estimate large-scale and small-scale errors, respectively (left). Assimilation of sea level (twin experiment) results in non-trivial improvements in velocity (right); Results of PS (red) are nearly indistinguishable from a global fine grid smoother (blue). Dashed lines are corresponding formal error estimates. (The green curve is a global coarse grid smoother based on (a) alone.)*

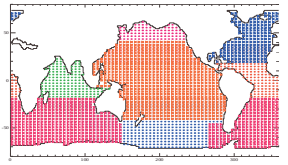

Fig 6: Partitioning Employed in the ECCO-2 PKF/PS [2]. *The global domain is decomposed into eight partitions that consist of seven overlapping regional cells and a single global cell. Model state error and process noise (control modeled as wind error) are estimated separately within each partition. Each partitioned filter/smoother is itself a reduced-state filter/smoother based on coarse-grid vertical dynamic mode decomposition. Regional state reductions consist of the first 5 baroclinic modes on a coarse grid (5° × 3°). The global cell consists of barotropic velocity and corresponding sea level error estimates on a 6° × 6° grid.*

Conclusion

Understanding the nature of model errors and establishing sensible models (approximations) of these uncertainties are central to the consistency and computational efficacy of data assimilation.

Temporal evolution of data assimilated state estimates ($\mathbf{x}_t$ in Eq 1) are physically inconsistent when the corresponding source of the model errors (process noise, system error, etc; $\mathbf{u}_t$ in Eq 1) are not determined. Inconsistencies, such as physical imbalances, can result in spurious inferences. The smoother coherently solves both errors of the state and errors of the model (Eq 8). Such estimation requires explicit modeling of model error sources by specific physical and numerical processes (e.g., forcing error, finite difference error, etc; $\mathbf{G}$ in Eq 1).

The partitioned Kalman filter (PKF) and partitioned smoother (PS) provide practical means for data assimilation using global state-of-the-art resolution models. Identifying sensible partitioning of model errors (Eq 10; e.g., Figs 5 & 6) is key to the efficacy of PKF/PS.

References

[1] Fukumori, I., A partitioned Kalman filter and smoother, *Mon. Weather Rev.*, **130**, 1370-1383, 2002.

[2] Fukumori, I., *et al.*, A prototype GODAE routine global ocean data assimilation system: ECCO-2, *"En route to GODAE"* (this symposium).

[3] Lee, T., *et al.*, Upper-ocean heat budget inferred from ECCO-2 ocean data assimilation system, *"En route to GODAE"* (this symposium).

[4] McKinley, G., *Interannual Variability of Air-Sea Fluxes of Carbon Dioxide and Oxygen*, Ph.D. thesis, MIT, Cambridge, Mass., 169pp, 2002.

[5] Menemenlis, D., *et al.*, Calibrating the ECCO ocean general circulation model using Green's functions, *"En route to GODAE"* (this symposium).

[6] Stammer, D., *et al.*, Variations in the large-scale ocean circulation estimated by combining satellite and WOCE data with a numerical model, (this symposium).